

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Behan, Dominic P.
Chalmers, Derek T.
Liaw, Chen
Lin, I-Lin
Lowitz, Kevin P.
Chen, Ruoping

(ii) TITLE OF INVENTION: Endogenous, Constitutively Activated
G Protein-Coupled Orphan Receptors

(iii) NUMBER OF SEQUENCES: 45

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE:
(B) STREET:
(C) CITY:
(D) STATE:
(E) COUNTRY:
(F) ZIP:

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Michael P. Straher
- (B) REGISTRATION NUMBER: 38,325

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE:
- (B) TELEFAX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGAGGATCCATGGCCTGGTTCTCAGC

26

(3) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CACAAGCTTAGRCCRTCCMGR CARTTCCA

29

(4) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGAGAAGCTTCTGGCGGCGATGAACGCTAG

30

(5) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ACAGGATCCAGGTGGCTGCTAGCAAGAG

28

(6) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTTAAGCTTAAATGAACGAAGACCCGAAG

30

(7) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGAGGATCCCCAGAGCATCACTAGCAT

27

(8) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 530 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGAGGATCCA TGGCCTGGTT CTCAGCCGGC TCAGGCAGTG TGAATGTGAG CATAGACCCA

60

GCAGAGGAAC CTACAGGCCC AGCTACACTG CTGCCCTCTC CCAGGGCCTG GGATGTGGTG

120

CTGTGCATCT CAGGCACCCT GGTGTCCTGC GAGAATGCTC TGGTGATGGC CATCATTTGTG

180

GGCAGCCTG CCTTCCGCGC CCCCATGTTT CTGCTGGTGG GCAGCTTGGC CGTAGCAGAC 240
CTGCTGGCAG GCCTGGGCCT GGTCTGCAC TTCGCTGCTG ACTTCTGTAT TGGCTCACCA 300
GAGATGAGCT TGGTGCTGGT TGGCGTGCTA GCAACGGCCT TTACTGCCAG CATCGGCAGC 360
CTGCTGGCCA TCACCGTTGA CCGCTACCTT TCCCTGTACA ACGCCCTCAC CTACTACTCA 420
GAGACAACAG TAACTCGAAC CTACGTGATG CTGGCCTTGG TGTGGGTGGG TGCCCTGGGC 480
CTGGGGCTGG TTCCCGTGCT GGCCTGGAAC TGCCGGGACG GTCTAAGCTT 530

(9) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 601 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AAGCTTCTGG CGGCGATGAA CGCTAGCGCC GCCGCGCTCA ACGAGTCCCA GGTGGTGGCA 60
GTAGCGGCCG AGGGAGCGGC AGCTGCGGCT ACAGCAGCAG GGACACCGGA CACCAGCGAA 120
TGGGGACCTC CGGCAGCATC CGCGGCGCTG GGAGGCGGCG GAGGACCTAA CGGGTCACTG 180
GAGCTGTCTT CGCAGCTGCC CGCAGGACCC TCAGGACTTC TGCTTTCGGC AGTGAATCCC 240
TGGGATGTGC TGCTGTGCGT GTCGGGGACT GTGATCGCAG GCGAAAATGC GCTGGTGGTG 300
GCGCTCATCG CATCCACTCC CGCGCTGCGC ACGCCCATGT TTGTGCTCGT GGGTAGTCTG 360
GCCACTGCTG ACCTGCTGGC GGGCTGTGGC CTCATCCTAC ACTTCGTGTT CCAGTACGTG 420

GTGCCCTCGG AGACTGTGAG CCTGCTCATG GTGGGCTTCC TGGTGGCGTC CTCGCCGCC 480
TCAGTCAGCA GCCTGCTCGC TATCACAGTG GACCGTTACC TGTCCCTTTA CAACGCGCTC 540
ACCTACTACT CGCGCCGGAC CCTGTTGGGC GTGCACCTCT TGCTAGCAGC CACCTGGATC 600
C 601

(10) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 510 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AAGCTTAAAA TGAACGAAGA CCCGAAGGTC AATTTAAGCG GGCTGCCTCG GGACTGTATA 60
GAAGCTGGTA CTCCGGAGAA CATCTCAGCC GCTGTCCCCT CCCAGGGCTC TGTTGTGGAG 120
TCAGAACCCG AGCTCGTTGT CAACCCCTGG GACATTGTCT TGTGCAGCTC AGGAACCCTC 180
ATCTGCTGTG AAAATGCCGT CGTGGTCCTT ATCATCTTCC ACAGCCCCAG CCTGCGAGCA 240
CCCATGTTCC TGCTGATAGG CAGCCTGGCT CTTGCAGACC TGCTGGCTGG TCTGGGACTC 300
ATCATCAATT TTGTTTTTGC CTACCTGCTT CAGTCAGAAG CCACCAAGCT GGTCACAATT 360
GGACTCATTG TCGCCTCTTT CTCTGCCTCT GTCTGCAGTT TGCTGGCTAT CACTGTGGAC 420
CGCTACCTCT CGCTGTATTA CGCCCTGACG TACCACTCCG AGAGGACCGT CACCTTTACC 480

TATGTCATGC TAGTGATGCT CTGGGGATCC

510

(11) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CTTAAGCTTGTGGCATTGTTACT

24

(12) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TCTGGATCCTTGGCCAGGCAGTGGAAGT

28

(13) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GAGAATTCAC TCCTGAGCTC AAGATGAACT

30

(14) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGGGATCCCC GTAAGTGGAG CACTTCAGAT

30

(15) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1050 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ATGAACTCCA CCTTGGATGG TAATCAGAGC AGCCACCCTT TTTGCCTCTT GGCATTTGGC	60
TATTTGGAAA CTGTCAATTT TTGCCTTTTG GAAGTATTGA TTATTGTCTT TCTAACTGTA	120
TTGATTATTT CTGGCAACAT CATTGTGATT TTTGTATTTC ACTGTGCACC TTTGTTGAAC	180
CATCACACTA CAAGTTATTT TATCCAGACT ATGGCATATG CTGACCTTTT TGTGTTGGGTG	240
AGCTGCGTGG TCCCTTCTTT ATCACTCCTC CATCACCCCC TTCCAGTAGA GGAGTCCTTG	300
ACTTGCCAGA TATTTGGTTT TGTAAGTATCA GTTCTGAAGA GCGTCTCCAT GGCTTCTCTG	360
GCCTGTATCA GCATTGATAG ATACATTGCC ATTACTAAAC CTTTAACCTA TAATACTCTG	420
GTTACACCCT GGAGACTACG CCTGTGTATT TTCCTGATTT GGCTATACTC GACCCTGGTC	480
TTCTGCCTT CTTTTTCCA CTGGGGCAA CCTGGATATC ATGGAGATGT GTTTCAGTGG	540
TGTGCGGAGT CCTGGCACAC CGACTCCTAC TTCACCCTGT TCATCGTGAT GATGTTATAT	600
CCCCCAGCAG CCCTTATTGT CTGCTTCACC TATTTCAACA TCTTCCGCAT CTGCCAACAG	660
CACACAAAGG ATATCAGCGA AAGGCAAGCC CGCTTCAGCA GCCAGAGTGG GGAGACTGGG	720
GAAGTGCAGG CCTGTCCTGA TAAGCGCTAT GCCATGGTCC TGTTTCGAAT CACTAGTGTA	780
TTTACATCC TCTGGTTGCC ATATATCATC TACTTCTTGT TGGAAAGCTC CACTGGCCAC	840
AGCAACCGCT TCGCATCCTT CTTGACCACC TGGCTTGCTA TTAGTAACAG TTTCTGCAAC	900
TGTGTAATTT ATAGTCTCTC CAACAGTGTA TTCAAAGAG GACTAAAGCG CCTCTCAGGG	960
GCTATGTGTA CTTCTTGTC AAGTCAGACT ACAGCCAACG ACCCTTACAC AGTTAGAAGC	1020
AAAGGCCCTC TTAATGGATG TCATATCTGA	1050

(16) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 349 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```

Met Asn Ser Thr Leu Asp Gly Asn Gln Ser Ser His Pro Phe Cys Leu
1           5           10           15

Leu Ala Phe Gly Tyr Leu Glu Thr Val Asn Phe Cys Leu Leu Glu Val
           20           25           30

Leu Ile Ile Val Phe Leu Thr Val Leu Ile Ile Ser Gly Asn Ile Ile
           35           40           45

Val Ile Phe Val Phe His Cys Ala Pro Leu Leu Asn His His Thr Thr
           50           55           60

Ser Tyr Phe Ile Gln Thr Met Ala Tyr Ala Asp Leu Phe Val Gly Val
65           70           75           80

Ser Cys Val Val Pro Ser Leu Ser Leu Leu His His Pro Leu Pro Val
           85           90           95

Glu Glu Ser Leu Thr Cys Gln Ile Phe Gly Phe Val Val Ser Val Leu
           100          105          110

Lys Ser Val Ser Met Ala Ser Leu Ala Cys Ile Ser Ile Asp Arg Tyr
           115          120          125

```

Ile Ala Ile Thr Lys Pro Leu Thr Tyr Asn Thr Leu Val Thr Pro Trp
130 135 140

Arg Leu Arg Leu Cys Ile Phe Leu Ile Trp Leu Tyr Ser Thr Leu Val
145 150 155 160

Phe Leu Pro Ser Phe Phe His Trp Gly Lys Pro Gly Tyr His Gly Asp
165 170 175

Val Phe Gln Trp Cys Ala Glu Ser Trp His Thr Asp Ser Tyr Phe Thr
180 185 190

Leu Phe Ile Val Met Met Leu Tyr Ala Pro Ala Ala Leu Ile Val Cys
195 200 205

Phe Thr Tyr Phe Asn Ile Phe Arg Ile Cys Gln Gln His Thr Lys Asp
210 215 220

Ile Ser Glu Arg Gln Ala Arg Phe Ser Ser Gln Ser Gly Glu Thr Gly
225 230 235 240

Glu Val Gln Ala Cys Pro Asp Lys Arg Tyr Ala Met Val Leu Phe Arg
245 250 255

Ile Thr Ser Val Phe Tyr Ile Leu Trp Leu Pro Tyr Ile Ile Tyr Phe
260 265 270

Leu Leu Glu Ser Ser Thr Gly His Ser Asn Arg Phe Ala Ser Phe Leu
275 280 285

Thr Thr Trp Leu Ala Ile Ser Asn Ser Phe Cys Asn Cys Val Ile Tyr
290 295 300

Ser Leu Ser Asn Ser Val Phe Gln Arg Gly Leu Lys Arg Leu Ser Gly
305 310 315 320

Ala Met Cys Thr Ser Cys Ala Ser Gln Thr Thr Ala Asn Asp Pro Tyr

325

330

335

Thr Val Arg Ser Lys Gly Pro Leu Asn Gly Cys His Ile

340

345

(17) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

AGGAAGCTTT AAATTTCCAA GCCATGAATG

30

(18) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ACCGAATTCA GATTACATTT GATTACTAT G

31

(19) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1086 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ATGAATGAAT CCAGGTGGAC TGAATGGAGG ATCCTGAACA TGAGCAGTGG CATTGTGAAT	60
GTGTCCGAGC GTCACCTCTG CCCACTTGGA TTTGGCCACT ACAGTGTGGT GGATGTCTGC	120
ATCTTCGAGA CAGTGGTTAT TGTGTTGCTG ACATTTCTGA TCATTGCTGG GAATCTAACA	180
GTTATCTTTG TCTTTCATTG TGCTCCACTG TTACATCATT ATACTACCAG CTATTTTCATT	240
CAGACGATGG CATATGCTGA TCTTTTCGTT GGAGTTAGCT GCTTGTTTCC TACTCTGTCA	300
CTTCTCCACT ACTCCACAGG TGTCCACGAG TCATTGACTT GCCAGGTTTT TGGATATATC	360
ATCTCAGTTC TAAAAAGTGT TTCTATGGCA TGTCTTGCTT GCATCAGTGT GGATCGTTAT	420
CTTGCAATAA CCAAGCCTCT TTCCTACAAT CAACTGGTCA CCCCTTGTCG CTTGAGAATT	480
TGCATTATTT TGATCTGGAT CTACTCCTGC CTAATTTTCT TGCCTTCCTT TTTTGGCTGG	540
GGGAAACCTG GTTACCATGG TGACATTTTT GAATGGTGTG CCACGTCTTG GCTCACCAGT	600
GCCTATTTTA CTGGCTTTAT TGTTTGTTTA CTTTATGCTC CTGCTGCCTT TGTGTCTGC	660
TTCACTTACT TCCACATTTT CAAAATTGTC CGTCAGCACA CCAAAGAGAT AAATGACCGA	720
AGAGCCCGAT TCCCTAGTCA TGAGGTAGAT TCTTCCAGAG AGACTGGACA CAGCCCTGAC	780

CGTCGCTACG CCATGGTTTT GTTTAGGATA ACCAGTGTAT TTTATATGCT GTGGCTCCCC 840

TATATAATTT ACTTTCTTCT AGAAAGCTCC CGGGTCTTGG ACAATCCAAC TCTGTCCTTC 900

TTAACAACCT GGCTTGCAAT AAGTAATAGT TTTTGTA ACT GTGTAATATA CAGCCTCTCC 960

AACAGCGTTT TCCGGCTAGG CCTCCGAAGA CTGTCTGAGA CAATGTGCAC ATCCTGTATG 1020

TGTGTGAAGG ATCAGGAAGC ACAAGAACCC AAACCTAGGA AACGGGCTAA TTCTTGCTCC 1080

ATTTGA 1086

(20) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 361 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met	Asn	Glu	Ser	Arg	Trp	Thr	Glu	Trp	Arg	Ile	Leu	Asn	Met	Ser	Ser
1				5					10					15	
Gly	Ile	Val	Asn	Val	Ser	Glu	Arg	His	Ser	Cys	Pro	Leu	Gly	Phe	Gly
			20					25					30		
His	Tyr	Ser	Val	Val	Asp	Val	Cys	Ile	Phe	Glu	Thr	Val	Val	Ile	Val
			35					40					45		
Leu	Leu	Thr	Phe	Leu	Ile	Ile	Ala	Gly	Asn	Leu	Thr	Val	Ile	Phe	Val
			50					55					60		

Phe His Cys Ala Pro Leu Leu His His Tyr Thr Thr Ser Tyr Phe Ile
65 70 75 80

Gln Thr Met Ala Tyr Ala Asp Leu Phe Val Gly Val Ser Cys Leu Val
85 90 95

Pro Thr Leu Ser Leu Leu His Tyr Ser Thr Gly Val His Glu Ser Leu
100 105 110

Thr Cys Gln Val Phe Gly Tyr Ile Ile Ser Val Leu Lys Ser Val Ser
115 120 125

Met Ala Cys Leu Ala Cys Ile Ser Val Asp Arg Tyr Leu Ala Ile Thr
130 135 140

Lys Pro Leu Ser Tyr Asn Gln Leu Val Thr Pro Cys Arg Leu Arg Ile
145 150 155 160

Cys Ile Ile Leu Ile Trp Ile Tyr Ser Cys Leu Ile Phe Leu Pro Ser
165 170 175

Phe Phe Gly Trp Gly Lys Pro Gly Tyr His Gly Asp Ile Phe Glu Trp
180 185 190

Cys Ala Thr Ser Trp Leu Thr Ser Ala Tyr Phe Thr Gly Phe Ile Val
195 200 205

Cys Leu Leu Tyr Ala Pro Ala Ala Phe Val Val Cys Phe Thr Tyr Phe
210 215 220

His Ile Phe Lys Ile Cys Arg Gln His Thr Lys Glu Ile Asn Asp Arg
225 230 235 240

Arg Ala Arg Phe Pro Ser His Glu Val Asp Ser Ser Arg Glu Thr Gly
245 250 255

His Ser Pro Asp Arg Arg Tyr Ala Met Val Leu Phe Arg Ile Thr Ser
260 265 270

Val Phe Tyr Met Leu Trp Leu Pro Tyr Ile Ile Tyr Phe Leu Leu Glu
275 280 285

Ser Ser Arg Val Leu Asp Asn Pro Thr Leu Ser Phe Leu Thr Thr Trp
290 295 300

Leu Ala Ile Ser Asn Ser Phe Cys Asn Cys Val Ile Tyr Ser Leu Ser
305 310 315 320

Asn Ser Val Phe Arg Leu Gly Leu Arg Arg Leu Ser Glu Thr Met Cys
325 330 335

Thr Ser Cys Met Cys Val Lys Asp Gln Glu Ala Gln Glu Pro Lys Pro
340 345 350

Arg Lys Arg Ala Asn Ser Cys Ser Ile
355 360

(21) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

AGCGAATTCT GCCCACCCTA CGCCGAGGTG CT

32

(22) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TGCGGATCCG CCAGCTCTTG AGCCTGCACA

30

(23) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1381 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GGCCTTATCT TTCCAGTCGT CCAGCATGCT CTGCCCACCC CAGCCGAGG TGCACTGACC	60
ATGAGCCTCA ACTCCTCCCT CAGCTGCAGG AAGGAGCTGA GTAATCTCAC TGAGGAGGAG	120
GGTGGCGAAG GGGGCGTCAT CATCACCAG TTCATCGCCA TCATTGTCAT CACCATTTTT	180
GTCTGCCTGG GAAACCTGGT CATCGTGGTC ACCTTGACCA AGAAGTCCTA CCTCCTCACC	240
CTCAGCAACA AGTTCGTCTT CAGCCTGACT CTGTCCAAC TCTGCTGTC CGTGTTGGTG	300
CTGCCTTTTG TGGTGACGAG CTCCATCCGC AGGGAATGGA TCTTTGGTGT AGTGTGGTGC	360

AACTTCTCTG CCCTCCTCTA CCTGCTGATC AGCTCTGCCA GCATGCTAAC CCTCGGGGTC 420

ATTGCCATCG ACCGCTACTA TGCTGTCCTG TACCCCATGG TGTACCCCAT GAAGATCACA 480

GGGAACCGGG CTGTGATGGC ACTTGTCTAC ATCTGGCTTC ACTCGCTCAT CGGCTGCCTG 540

CCACCCCTGT TTGGTTGGTC ATCCGTGGAG TTTGACGAGT TCAAATGGAT GTGTGTGGCT 600

GCTTGGCACC GGGAGCCTGG CTACACGGCC TTCTGGCAGA TCTGGTGTGC CCTCTTCCCC 660

TTTCTGGTCA TGCTGGTGTG CTATGGCTTC ATCTTCCGCG TGGCCAGGGT CAAGGCACGC 720

AAGGTGCACT GTGGCACAGT CGTCATCGTG GAGGAGGATG CTCAGAGGAC CGGGAGGAAG 780

AACTCCAGCA CCTCCACCTC CTCTTCAGGC AGCAGGAGGA ATGCCTTTCA GGGTGTGGTC 840

TACTCGGCCA ACCAGTGCAA AGCCCTCATC ACCATCCTGG TGGTCCTCGG TGCCTTCATG 900

GTACACCTGGG GCCCCTACAT GGTGTGCATC GCCTCTGAGG CCCTCTGGGG GAAAAGCTCC 960

GTCTCCCCGA GCCTGGAGAC TTGGGCCACA TGGCTGTCCT TTGCCAGCGC TGTCTGCCAC 1020

CCCCTGATCT ATGGACTCTG GAACAAGACA GTTCGCAAAG AACTACTGGG CATGTGCTTT 1080

GGGGACCGGT ATTATCGGGA ACCATTTGTG CAACGACAGA GGACTTCCAG GCTCTTCAGC 1140

ATTTCCAACA GGATCACAGA CCTGGGCCTG TCCCCACACC TCACTGCGCT CATGGCAGGT 1200

GGACAGCCCC TGGGGCACAG CAGCAGCACG GGGGACACTG GCTTCAGCTG CTCCCAGGAC 1260

TCAGGTAACC TGCGTGCTTT ATAAGCCTCT CACCTGTCGC GTTTTCCCTG TGTTCGCTTT 1320

CCCCCGTGTC GCGTTTCCCC TGTGCAGGCT CAAGAGCTGG CGGAGGGGCA TTTCCCACGG 1380

TG 1382

(24) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 407 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

MSLNSSLSCR KELSNTLEE GGEGGVIITQ FIAIIVITIF VCLGNLVIV
TLYKKSyllt LSNKFVFSlt LSNFLLSVLV LPFVVTSSIR REWIFGVVWC

NFSALLYLLI SSASMLTLGV IAIDRYYAVL YPMVYPMKIT GNRAVMALVY
IWLHSLIGCL PPLFGWSSVE FDEFKWCVA AWHREPGYTA FWQIWCAFP
FLVMLVCYGF IFRVARVKAR KVHCGTVVIV EEDAQRTGRK NSSTSTSSSG
SRRNAFQGVV YSANQCKALI TILVVLGAFM VTWGPYMVVI ASEALWGKSS
VSPSLETWAT WLSFASAVCH PLIYGLWNKT VRKELLMCF GDRYYREPFV
QRQRTSRLFS ISNRITDLGL SPHLTALMAG GQPLGHSSST GDTGFSCSQD

SGNLRAL

(25) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GGAAGCTTCA GGCCCAAAGA TGGGGAACAT

30

(26) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GTGGATCCAC CCGCGGAGGA CCCAGGCTAG

30

(27) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1697 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

ACTCCCAAAG TGCTGGGCTT ACAGGTGTAA GCCATCATGT CCAGCCGTTT AGATATTCTA 60
GTTGAATTGG AGTTGGTGGG CTAGTACACC TTCTAAATTA AATGAGTAAA GGATTTAGAA 120

TGGTGCCTGA CACACAGTAG GTGCTACATT CATGTTAGCT ACTATTATAA ACCTTTCCTG 180

CCTCTGACTT TCAGGGTCTT GCCCACCACC AGCGATGCCC AGCCCTTGGT AGAGCTTGAA 240

CCACCTTCTA TAAACAGGAT GCGGTGGAG AGACAGGCC AGTCCCTGAG CCCATGAGGA 300

GTGTGGCCCC TTCAGGCCA AAGATGGGA ACATCACTGC AGACAACTCC TCGATGAGCT 360

GTACCATCGA CCATACCATC CACCAGACGC TGGCCCCGGT GGTCTATGTT ACCGTGCTGG 420

TGGTGGGCTT CCCGGCCAAC TGCCTGTCCC TCTACTTCGG CTACCTGCAG ATCAAGGCC 480

GGAACGAGCT GGGCGTGAC CTGTGCAACC TGACGGTGGC CGACCTCTC TACATCTGCT 540

CGCTGCCCTT CTGGCTGCAG TACGTGCTGC AGCAGACAA CTGGTCTCAC GCGACCTGT 600

CCTGCCAGGT GTGCGGCATC CTCCTGTACG AGAACATCTA CATCAGCGTG GGCTTCCTCT 660

GCTGCATCTC CGTGGACCGC TACCTGGCTG TGGCCCATCC CTTCCGCTTC CACCAGTTCC 720

GGACCTGAA GCGGCCCGTC GCGTCAGCG TGGTCATCTG GGCCAAGGAG CTGCTGACCA 780

GCATCTACTT CCTGATGCAC GAGGAGGTCA TCGAGGACGA GAACCAGCAC CGCGTGTGCT 840

TTGAGCACTA CCCCATCCAG GCATGGCAGC GCGCCATCAA CTACTACCGC TTCCTGGTGG 900

GCTTCCTCTT CCCCATCTGC CTGCTGCTGG CGTCCTACCA GGCATCCTG CGCGCCGTGC 960

GCCGGAGCCA CGGCACCCAG AAGAGCCGCA AGGACCAGAT CCAGCGGCTG GTGCTCAGCA 1020

CCGTGGTCAT CTTCTGGCC TGCTTCCTGC CCTACCACGT GTTGCTGCTG GTGCGCAGCG 1080

TCTGGGAGGC CAGCTGCGAC TTCGCCAAGG GCGTTTCAA CGCCTACCAC TTCTCCCTCC 1140

TGCTCACCAG CTTCAACTGC GTCGCCGACC CCGTGCTCTA CTGCTTCGTC AGCGAGACCA 1200

CCCACCGGGA CCTGGCCCGC CTCCGCGGGG CCTGCCTGGC CTTCTCACC TGCTCCAGGA 1260

CCGGCCGGGC CAGGGAGGCC TACCCGCTGG GTGCCCCGA GGCCTCCGGG AAAAGCGGGG 1320

CCCAGGGTGA GGAGCCCGAG CTGTTGACCA AGCTCCACCC GGCCTTCCAG ACCCCTAACT 1380

CGCCAGGGTC GGGCGGGTTC CCCACGGGCA GGTTCGCCTA GCCTGGGTCC TCCGCGGGTG 1440

GCTCCACGTG AGGCCTGAGC CTTAGCCCA CGGGCCTCAG GGCCTGCCGC CTCCTGCTTC 1500

CCTCGCTGCG GAGGCAGGGA AGCCCTGTA ACTCCGAAG CCTGCTCTCG CTTGCTGAGC 1560

CCGCTGGGAC CGCCGAGGCT GGAATAAGC CCCGGTTGGC TCGTGGGAAT AAGCCGTGTC 1620

CTCTGCCGCG GCTGCGATGT GGCCACGCTG GGGCTGCTGG TCGGGGAAA ACAGTGAAC 1680

GCGTCCCCTG GCCTGCT 1697

(28) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 365 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

MGNITADNSS MSCTIDHTIH QTLAPVVYVT VLVVGFPANC LSLYFGYLQI

KARNELGVYL CNLTVADLFY ICSLPFWLQY VLQHDNWSHG DLSCQVCGIL

LYENIYISVG FLCCISVDYR LAVAHPPFRFH QFRTLKAAVG VSVVIWAKEL
LTSIYFLMHE EVIEDENQHR VCFEHYPIQA WQRAINYYRF LVGFLFPICL
LLASYQGILR AVRRSHGTQK SRKDQIQRLV LSTVVIFLAC FLPYHVLLLV
RSVWEASCDF AKGVFNAYHF SLLLTSEFNCV ADPVLYCFVS ETTHRDRLRL
RGACLAFLTC SRTGRAREAY PLGAPEASGK SGAQGEEPEL LTKLHPAFQT
PNSPGSGGFP TGRLA

(29) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CTGGTCCTGC ACTTTGCTGC

20

(30) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

AGCATCACAT AGGTCCGTGT CAC

23

(31) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

ACCAGAAAGG GTGTGGGTAC ACTG

24

(32) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GGAACGAAAG GGCACCTTGG

20

(33) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GCTGCCTCGG GATTATTTAG

20

(34) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GCCTATTAGC AGGAACATGGGTG

23

(35) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GCTAGCGTTC ATCGCCGC

18

(36) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CTGGACTGTA TCGCCCCG

18

(37) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GATCTCTAGA ATGATGTGGG GTGCAGGCAG CC

32

(38) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CTAGGGTACC CGGACATCAC TGGGGGAGCG GGATC

35

(39) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GATCTCTAGA ATGCAGGGTG CAAATCCGGC C

31

(40) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 35 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

CTAGGGTACC CGGACCTCGC TGGGAGACCT GGAAC

35

(41) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

ATGTGGAACG CGACGCCCAG CG

22

(42) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

TCATGTATTA ATACTAGATT CT

42

(43) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

TACCATGTGG AACGCGACGC CCAGCGAAGA GCCGGGGT

38

(44) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

CGGAATTCAT GTATTAATAC TAGATTCTGT CCAGGCCCG

39

(45) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1101 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

ATGTGGAACG CGACGCCCAG CGAAGAGCCG GGGTTCAACC TCACACTGGC CGACCTGGAC

60

TGGGATGCTT CCCCCGGCAA CGACTCGCTG GGCGACGAGC TGCTGCAGCT CTTCCCCGCG

120

CCGCTGCTGG CGGGCGTCAC AGCCACCTGC GTGGCACTCT TCGTGGTGGG TATCGCTGGC

180

AACCTGCTCA CCATGCTGGT GGTGTCGCGC TTCCGCGAGC TGCGCACCAC CACCAACCTC

240

TACCTGTCCA GCATGGCCTT CTCCGATCTG CTCATCTTCC TCTGCATGCC CCTGGACCTC	300
GTTCGCCTCT GGCAGTACCG GCCCTGGAAC TTCGGCGACC TCCTCTGCAA ACTCTTCCAA	360
TTCGTCAGTG AGAGCTGCAC CTACGCCACG GTGCTCACCA TCACAGCGCT GAGCGTCGAG	420
CGCTACTTCG CCATCTGCTT CCCACTCCGG GCCAAGGTGG TGGTCACCAA GGGGCGGGTG	480
AAGCTGGTCA TCTTCGTCAT CTGGGCCGTG GCCTTCTGCA GCGCCGGGCC CATCTTCGTG	540
CTAGTCGGGG TGGAGCACGA GAACGGCACC GACCCTTGGG ACACCAACGA GTGCCGCCCC	600
ACCGAGTTTG CGGTGCGCTC TGGACTGCTC ACGGTCATGG TGTGGGTGTC CAGCATCTTC	660
TTCTTCCTTC CTGTCTTCTG TCTCACGGTC CTCTACAGTC TCATCGGCAG GAAGCTGTGG	720
CGGAGGAGGC GCGGCGATGC TGTCGTGGGT GCCTCGCTCA GGGACCAGAA CCACAAGCAA	780
ACCGTGAAAA TGCTGGCTGT AGTGGTGTTT GCCTTCATCC TCTGCTGGCT CCCCTTCCAC	840
GTAGGGCGAT ATTTATTTTC CAAATCCTTT GAGCCTGGCT CCTTGGAGAT TGCTCAGATC	900
AGCCAGTACT GCAACCTCGT GTCCTTTGTC CTCTTCTACC TCAGTGCTGC CATCAACCCC	960
ATTCTGTACA ACATCATGTC CAAGAAGTAC CGGGTGGCAG TGTTCAGACT TCTGGGATTC	1020
GAACCCTTCT CCCAGAGAAA GCTCTCCACT CTGAAAGATG AAAGTTCTCG GGCCTGGACA	1080
GAATCTAGTA TTAATACATG A	1101

(46) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 366 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

MWNATPSEEP GFNLTLADLD WDASPGNDSL GDELLQLFPA PLLAGVTATC VALFVVGIAG
NLLTMLVVSF FRELRTTNL YLSSMAFSDL LIFLCMPDL VRLWQYRPWN FGDLLCKLFQ
FVSESTYAT VLTITALSVE RYFAICFPLR AKVVVTKGRV KLVIFVIWAV AFCSAGPIFV
LVGVEHENG TDPWDTNECRP TEFAVRSGLL TVMVVWSSIF FFLPVFCLTV LYSLIGRKLW
RRRRGDAVVG ASLRDQNHKQ TVKMLAVVVF AFILCWLPHF VGRYLFSSSF EPGSLEIAQI
SQYCNLVSFV LFYLSAAINP ILYNIMSKKY RVAVFRLLGF EPFSQRKLST LKDESSRAWT
ESSINT